**Review**

Yuzhuo Wang, Meng Zhu, Hongxia Ma and Hongbing Shen*

# Polygenic risk scores: the future of cancer risk prediction, screening, and precision prevention

**Abstract:** Genome-wide association studies (GWASs) have shown that the genetic architecture of cancers are highly polygenic and enabled researchers to identify genetic risk loci for cancers. The genetic variants associated with a cancer can be combined into a polygenic risk score (PRS), which captures part of an individual's genetic susceptibility to cancer. Recently, PRSs have been widely used in cancer risk prediction and are shown to be capable of identifying groups of individuals who could benefit from the knowledge of their probabilistic susceptibility to cancer, which leads to an increased interest in understanding the potential utility of PRSs that might further refine the assessment and management of cancer risk. In this context, we provide an overview of the major discoveries from cancer GWASs. We then review the methodologies used for PRS construction, and describe steps for the development and evaluation of risk prediction models that include PRS and/or conventional risk factors. Potential utility of PRSs in cancer risk prediction, screening, and precision prevention are illustrated. Challenges and practical considerations relevant to the implementation of PRSs in health care settings are discussed.

**Keywords:** cancer screening; genome-wide association study (GWAS); polygenic risk score (PRS); precision prevention; risk prediction model.

*Corresponding author: Hongbing Shen**, Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211166, Jiangsu, China; Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing 211166, Jiangsu, China; and Research Units of Cohort Study on Cardiovascular Diseases and Cancers, Chinese Academy of Medical Sciences, Beijing 100010, China, Phone: +86 25 868 68437, E-mail: hbshen@njmu.edu.cn. https://orcid.org/0000-0002-2581-5906
**Yuzhuo Wang,** Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu, China; and Department of Medical Informatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing, Jiangsu, China
**Meng Zhu,** Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu, China; and Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, Jiangsu, China
**Hongxia Ma,** Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu, China; Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, Jiangsu, China; and Research Units of Cohort Study on Cardiovascular Diseases and Cancers, Chinese Academy of Medical Sciences, Beijing, China

## Introduction

Cancer ranks as a leading cause of death and a most important barrier to increasing life expectancy in the world [1]. There were an estimated 19.3 million cases and 10 million cancer deaths worldwide in 2020 [2]. Consequently, the development of public health strategies for cancer prevention is critically important. Using risk prediction models to estimate the probability or risk of developing cancer for an individual can help make clinical decisions about whether the individual warrants an intervention, such as early detection and prevention of the cancer [3, 4]. Moreover, when communicated and understood properly, it can be used to guide at-risk individuals on personal health management by adopting healthier lifestyle or behavior [4].

Classical risk prediction models for common cancers often incorporate basic demographic characteristics (e.g., age and sex), lifestyle factors or environmental exposures (e.g., smoking status, alcohol consumption, body mass index), clinical risk factors (e.g., medication history, laboratory-based biomarkers, and imaging features), inherited mutations leading to a moderate-to-high risk of cancer (e.g., *BRCA1/BRCA2* for breast and ovarian cancer), and family history [5–7]. Most of them do not include risk associated with common susceptibility variants. As genetic susceptibility variants are the earliest measurable contributor to the heritable risk of common cancers,

genetic profiling could be considered useful for improving the assessment and management of cancer risk [3, 4, 8].

Indeed, large-scale genome-wide association studies (GWASs) have identified hundreds of single-nucleotide polymorphisms (SNPs) associated with susceptibility to common cancers [9]. For any given cancer, single genetic variant exhibits only modest effect on the disease risk. However, the impact of their cumulative effect across the genome is considerable for many cancers [10–14]. Polygenic risk scores (PRSs) measuring this cumulative effect of genetic risk have recently been shown to be consistently associated with cancer risk and are capable of identifying a larger fraction of the population with risk equivalent to rare monogenic mutations, leading to an increased interest in understanding the potential clinical and public health utility of polygenic risk prediction that might further refine the assessment of cancer risk, thereby improving cancer prevention and early detection [15, 16].

In this review, we begin with an overview of the major discoveries from cancer GWASs. We then describe methodologies of leveraging genetic risk factors for PRS construction and review the steps for the development and evaluation of risk prediction models that include PRS and/ or conventional risk factors. We also illustrate the potential utility of PRS in cancer risk prediction, screening, and precision prevention. Some of the challenges and practical considerations relevant to the implementation of PRSs in clinical or public health settings are discussed.

# Genetic inheritance of cancer

Epidemiological studies provide strong evidence that inherited genetic factors play an important role in the etiology of cancer [17–19]. Almost all common cancers show some degree of familial aggregation, with first-degree relatives of patients having two- to four-fold increased risk for developing the same cancer [20]. Indeed, family-based linkage studies have identified many rare high-penetrant mutations in cancer susceptibility genes (CSGs), including *BRCA1* and *BRCA2* for breast and ovarian cancers [21, 22], *APC* and *AXIN2* for colorectal cancer (CRC) [23–27], and *EGFR* for non-small cell lung cancer [28]. To date, more than a hundred CSGs with high-penetrant mutations (at least 5% of mutation carriers develop cancer) that confer greater than two-fold relative risks of cancer have been identified [29].

Although family-based linkage studies have been successful in mapping genes associated with susceptibility

to many cancers, the findings of this approach are mainly limited to genes with highly penetrant variations [30]. These variations are rare in the population and they account for a small fraction of the familial risk of the respective cancer [31–35]. Together with the observation of familial cancer patients without highly penetrant variations in CSGs [36], and the fact that the majority of cancer cases do not occur in highly affected families, it is believed that a large fraction of genetic susceptibility to cancer results from the combined effects of many common genetic variants with small effect sizes [30]. Genetic-association studies provide an efficient approach for identifying common genetic variants that confer modest disease risks [30]. These studies are generally based on the "common disease, common variant" hypothesis [37]. Early association studies usually used candidate gene approach, which tested the association between genetic variants of genes that may be involved in carcinogenesis (i.e., candidate genes) and the cancer itself. However, candidate gene approach is limited by its reliance on existing knowledge for the biology of the disease being investigated [30]. Additionally, the numerous examples of associations that cannot be replicated have also led to skepticism about this approach [38].

# The advent of genome-wide association study

In the past two decades, the completion of the Human Genome Project [39], the deposition of millions of SNPs into public databases (e.g., through the International HapMap Project [40] and the 1000 Genomes Project [41]) [42], and the rapid developments in high-throughput genotyping technologies [43] have enabled population-based GWASs with increasing sample sizes. Through testing hundreds of thousands to millions of genetic variants across the genomes of many unrelated individuals, GWAS compare the frequency of genetic variations in a large sample of cancer patients and those in matched controls to identify genetic variants associated with susceptibility to cancer [44]. Genotyping can be performed using genome-wide SNP arrays combined with imputation or whole-genome sequencing (WGS) [45].

As GWAS test millions of genetic variants, it is important to account for the multiple testing issues to reduce false-positive associations. The most commonly used threshold to declare statistical evidence of association in GWAS is $p < 5.0 \times 10^{-8}$. This value corresponds to the Bonferroni correction of 0.05 divided by approximately one million independent tests [46].

## Cancer risk loci identified by GWAS

GWAS have been very successful in identifying risk loci for a vast number of common cancers, including lung, breast, ovarian, prostate, colorectal, gastric, renal, and bladder cancers. GWAS have also been reported for several hematological malignancies such as chronic lymphocytic leukemia, Hodgkin lymphoma, chronic myeloid leukemia, and diffuse large B cell lymphoma. Additionally, GWAS have identified common susceptibility variants for several pediatric solid cancers, including Wilms tumor and neuroblastoma. As of April 2021, more than 2,820 associations between genetic variants and cancer risk with genome-wide significance ($p<5.0 \times 10^{-8}$) have been reported by GWASs (see Supplementary Table 1, Supplementary Figure 1) [9]. Although, for many of these cancers, the number and scale of GWAS in Europeans far exceed those in non-European populations, risk loci specific to other ethnicities, such as Asian, Hispanic, and African American, have also been identified [9]. Performing GWAS in diverse populations can reveal novel ancestry-specific loci associated with cancer susceptibility [47].

Data from GWASs demonstrated that, for almost all cancers that have been investigated, the genetic underpinnings are highly polygenic, with SNPs in many genes contributing to the heritable risk in the population. Most of the cancer susceptibility loci identified to date are associated with modest increases in risk, with per allele odds ratios (ORs) generally less than 1.5, which means that the average proportion of variance explained by each variant is small [44]. However, as the susceptibility loci identified by GWAS have to pass a very stringent threshold of statistical significance, there are likely many more SNPs (e.g., with weaker effect sizes and/or with smaller allele frequencies) that do not meet this criteria but still contribute to the heritable risk of a given cancer [44]. Quantifying the heritability explained by both known and potential susceptibility SNPs is therefore informative with respective to the genetic architecture of cancers.

## Heritability explained in a population

Heritability refers to the proportion of phenotypic variation in a population that is attributed to genetic factors. In particular, narrow-sense heritability is the proportion of phenotypic variation due to additive genetic factors, whereas broad-sense heritability is the phenotypic variation due to all genetic factors (e.g., dominance and epistasis) [48]. Historically, family-based studies have long been used to estimate the heritability of cancer [18].

GWASs have facilitated estimation of how much of the total phenotypic variation in a population that is due to additive genetic factors tagged by genotyped and imputed SNPs. This quantification of "SNP heritability" can provide general insights into the genetic architecture of cancers. Statistical modeling of heritability is an active area of method development and is beyond the scope of this review; we refer readers to the emerging literatures [49–53]. Popular methods such as genome-wide complex trait analysis (GCTA) [54] have recently been used to analyze GWAS data and have shown that common genetic variations explain a substantial fraction of the heritable risk for many cancers, with estimates of 10% for breast cancer, 11% for prostate cancer, 15% for lung cancer, 20% for gastric cancer, and 16% for CRC in Chinese populations [55]. Heritability analysis based on GWASs of European populations yielded similar results, with estimates of 10% for estrogen receptor-negative breast cancer [56], 38% for prostate cancer [56], 21% for lung cancer [56], 25% for gastric cancer [56], and 17% for CRC [57]. The available evidence suggests that a considerable fraction of the heritable risk for many cancers is mediated by numerous common genetic variants distributed throughout the genome, most of which can be captured by genome-wide genotyping and/or imputation.

## Application of GWAS findings in human genome epidemiology

An important objective of cancer genetics is to translate research findings into clinical and public health practice. Despite the immense amount of time required [58], emerging evidence highlights diverse areas that GWAS findings can have clinical applications, including (but not limited to) predicting risk of cancer, facilitating disease classification and subtyping, informing drug development and drug toxicity, guiding clinicians on cancer prognosis, and treatment-related complications [44, 45]. This review mainly focuses on the application of GWAS findings to cancer risk prediction and precision prevention. Although most of the genetic variants identified by GWAS have small effects on cancer risk and each individual variant is incapable of effectively predicting disease risk, their combined effect, in the form of PRS, has the potential to identify a substantial fraction of the population who are at high risk of a certain cancer, thereby improving health outcomes through early detection, prevention, or treatment [45].

# Polygenic risk scores

Information from large-scale GWAS provides us with an opportunity to develop risk prediction models that incorporate risk SNPs. With accurate risk prediction models, we can better advise individuals on appropriate screening and prevention. A commonly used approach towards this aim is the development of PRS, which provides a quantitative measure of the genetic risk burden of the disease over a set of genetic variants. PRS has shown the potential to improve the efficiency of existing cancer screening programs [4, 44]. The use of PRS can also help stratify individuals into groups with significantly different risks of cancer to inform strategies to prevent or delay the onset of cancer (e.g., chemoprevention and lifestyle modifications) [3, 4]. In this section, we begin with an overview of current methodologies for PRS construction, followed by steps for the development and evaluation of risk prediction models that include PRS and/or conventional risk factors (Figure 1, Table 1).

## Calculation of PRS

PRS is calculated by summing the risk alleles corresponding to a phenotype of interest for each individual, weighted by the effect size estimate from an independent GWAS on the phenotype [3]. Determining which SNPs to include ("SNP selection") and the disease-associated weights to assign to the selected SNPs ("weight calculation") are two critical methodological aspects of PRS construction. Imprecision in any of these aspects can lead to a decrease in prediction accuracy [3, 59].

## SNP selection

SNP selection is critical because they constitute the building blocks of PRS. GWAS for most of the common cancers provide direct evidence of polygenic susceptibility [44], which means that the heritable risk could be mediated by numerous common genetic variants each with small effects. Under such polygenic architecture, selecting the true set of susceptibility variants for PRS construction is particularly challenging [3]. The simplest and most commonly used approach is to select SNPs based on the predefined criteria (e.g., those passing a p-value threshold in a given GWAS), and then weight the SNPs according to the corresponding estimated regression coefficients (e.g., log OR parameters from logistic

regression model). Many studies have investigated the risk prediction ability of the PRS that is constructed from independent SNPs reaching genome-wide significant threshold [60–62]. These studies, however, may omit a vast amount of true susceptibility SNPs with smaller effect sizes [3, 59]. The predictive power of PRS may be improved by including additional SNPs that are below the genome-wide significant threshold. This allows the inclusion of signals from more susceptibility SNPs with smaller effects at the cost of adding noise from false-positive associations [3, 59]. Practically, the optimal threshold can be determined based on the performance of the PRS in an independent study population, or using cross-validation techniques [3].

Another challenge for SNP selection is the presence of linkage disequilibrium (LD). In the presence of LD, both disease susceptibility SNPs and their correlated neighbors can satisfy the p-value threshold criteria for selection. Inclusion of SNPs in LD would decrease the predictive accuracy of PRS [63]. This problem can be addressed by the "clumping/pruning and thresholding" approach, which is implemented in two popular software packages, namely PLINK and PRSice [64, 65]. The clumping step sorts SNPs based on the strength of association statistics, and removes the SNPs that are correlated with the strongest signal within LD blocks. The thresholding step eliminates SNPs with association p-values larger than a predefined threshold. Typically, a stringent LD threshold (e.g., $r^2 < 0.05$) is needed in the clumping step to eliminate redundant effects caused by correlated SNPs [3]. However, stringent LD pruning can also exclude susceptibility variants that are in LD but contain independent association signals, thereby reducing the predictive accuracy of PRS.

Stepwise regression, which is usually implemented after the inclusion of a set of variants that satisfied a predefined p-value threshold, can also be used for SNP selection [16, 66]. Stepwise regression retains SNPs in a semi-automated manner by successively adding or removing variants based solely on the test statistics of their estimated coefficients. This approach has the disadvantage of ignoring prior knowledge of LD structure [59].

## Weight calculation

Another critical aspect for PRS construction is the calculation of SNP weights. The commonly used weight for each SNP is the log OR derived from an independent GWAS. An extension of this approach is polygenic hazard score (PHS), which is also calculated as a weighted sum of risk alleles
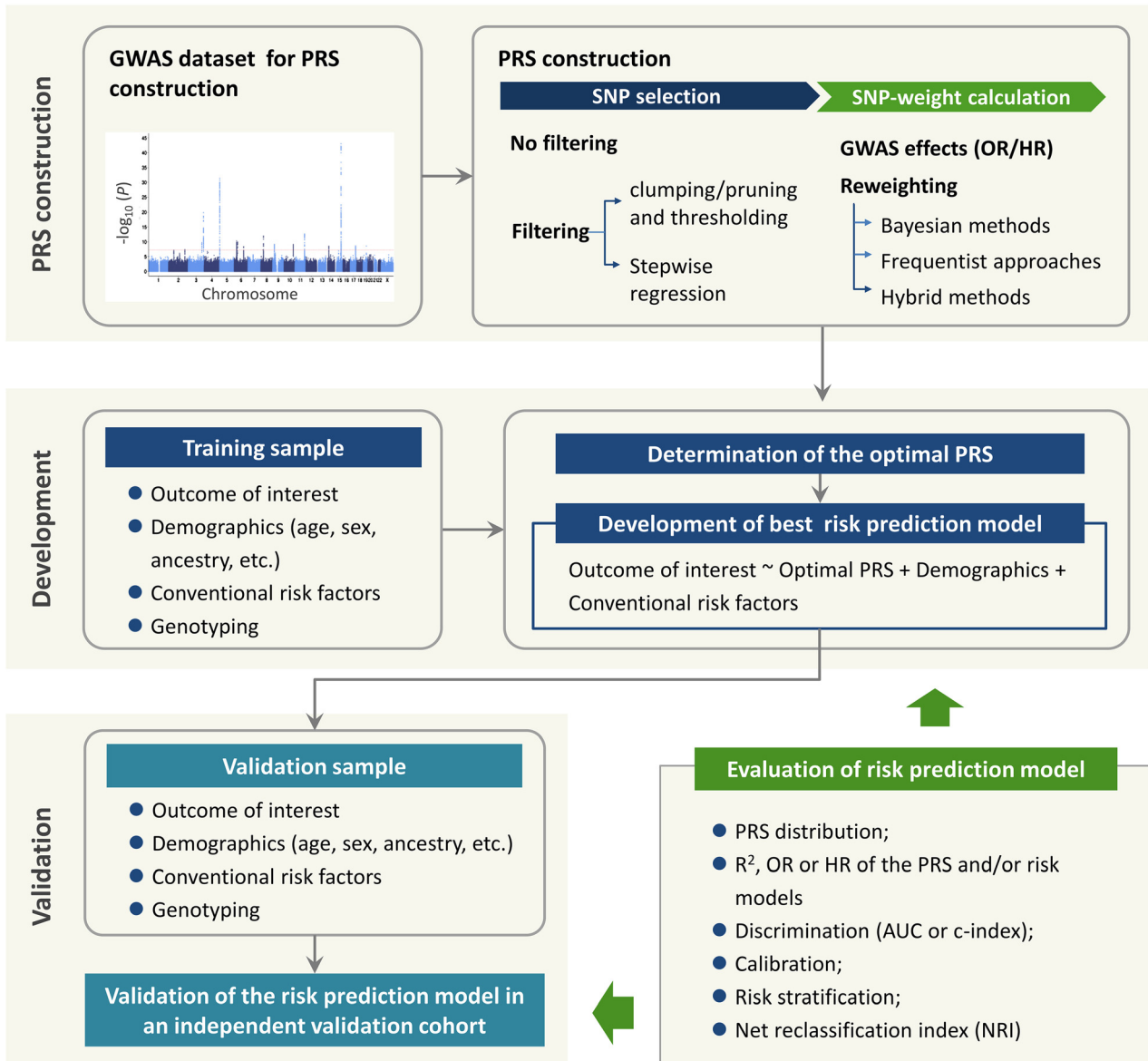
**Figure 1:** Development and validation of polygenic risk prediction models. The recommended steps for PRS construction, risk model development and validation are displayed. During PRS construction, genetic variants associated with an outcome of interest in a GWAS dataset are combined as a weighted sum of risk allele counts. Commonly used methods for "SNP selection" and "SNP-weight calculation" during the PRS construction procedure are shown. Performance of PRSs are evaluated in the training sample to select the optimal PRS. This optimal PRS is then added to a risk prediction model and may be combined with demographics (e.g., age, sex, and ancestry) and conventional risk factors (e.g., lifestyle factors or environmental exposures, clinical risk factors, inherited mutations leading to a moderate-to-high risk of cancer, and family history) to predict the outcome of interest. After model building procedure to select the best risk prediction model, this model is validated in an independent sample. For the evaluation of risk prediction model, the distribution of the PRS, the proportion of variance explained ($R^2$) and effect size estimates (e.g., ORs, HRs) of the PRS and/or risk models should be described. Performance of the risk prediction model in terms of discrimination, calibration, risk stratification, and NRI should also be assessed. Results from risk model evaluation should be reported for both the training and validation samples for comparison. PRS, polygenic risk score; GWAS, genome-wide association study; OR, odds ratio; HR, hazard ratio; $R^2$, the proportion of variance explained; AUC, area under the receiver operating characteristic curve; NRI, net reclassification index.

but in this case the weight for each variant is its log hazard ratio (HR) estimated from survival models [66–69]. Recently, a variety of Bayesian methods and Frequentist approaches have been proposed to optimize PRS performance by adjusting SNP weights [59].

LDpred uses Bayesian framework to adjust SNP weights from GWAS summary statistics by assuming a prior for the genetic architecture and using LD information from an external reference panel [70]. When applied to simulation data, LDpred outperforms the traditional

**Table 1:** Recommended steps for evaluating the performance of polygenic risk prediction models.

| Recommended steps | Description |
|---|---|
| PRS distribution | Distribution of the PRS (e.g., histogram or density plot of PRS distribution, mean, median, standard deviation, IQR, range, etc.). |
| Predictive ability | Proportion of variance explained ($R^2$) and effect size estimates (e.g., ORs or hazards ratios from regression models) used to evaluate the PRS and/or risk models. |
| Discrimination | Discrimination is a measure of how well the risk prediction model can separate those who will develop the disease in the future and those who will not. The discrimination ability is commonly quantified by AUC or the c-statistic. The AUC (or c-statistic) ranges from 50% to 100%. The greater the AUC (or c-statistic), the better the risk discriminatory ability of the model. |
| Calibration | Calibration reflects the ability of a risk prediction model to correctly estimate the risk for subjects with different risk factor profiles in the underlying population. Calibration of a risk prediction model can be evaluated through graphical representation of the relationship between predicted and observed risk, or by using statistical tests (e.g., the Hosmer–Lemeshow test). |
| Risk stratification | Risk stratification refers to the ability of the risk prediction model to separate the subjects into categories with sufficient distinct degrees of absolute risk to drive clinical or personal decisions. The risk stratification ability of a model can be quantified by (i) the proportions of individuals who are allocated into clinically relevant risk categories; and (ii) the proportion of patients who will develop a disease in the future that may be identified as being at high risk. |
| NRI | The NRI is commonly used to quantify whether new risk factors provides clinically relevant improvements in risk prediction. The widely recommended method is to calculate the category-based NRI, with the formulas shown as follows: <br> **Event NRI** = P (up\|event) − P (down\|event) = (number of persons with the event classified up − number of persons with the event classified down)/total number of persons with the event <br> (The net percentage of persons with the event correctly reclassified upward. It can be interpreted as a percentage with a range from −100 to 100%.) <br> **Nonevent NRI** = P (down\|nonevent) − P (up\|nonevent) = (number of persons without the event classified down − number of persons without the event classified up)/total number of persons without the event <br> (The net percentage of persons without the event correctly reclassified downward. It can be interpreted as a percentage with a range from −100 to 100%.) <br> **Overall NRI** = event NRI + nonevent NRI <br> (The sum of the net percentages of correctly reclassified persons with and without the event of interest. This statistic does not represent a percentage. Overall NRI can range from −2 to 2.) |

PRS, polygenic risk score; IQR, interquartile range; ORs, odds ratios; NRI, net reclassification index; AUC, area under the curve.

approach of "clumping/pruning and thresholding", particularly at large sample sizes [70]. AnnoPred improved LDpred by leveraging various types of genomic and epigenomic functional annotations to up-weight SNPs with higher likelihood of functionality [71]. Compared with LDpred, AnnoPred achieved higher prediction accuracy and better risk stratification ability [71].

Frequentist approaches, including linear mixed model (LMM) and penalized regression, have been utilized for PRS calculation [59]. Genetic Risk Scores Inference (GeRSI) is a method that utilizes LMM for PRS calculation. It includes SNPs below a certain p-value threshold as fixed effects and treats the rest of the SNPs as random effects within the framework of the widely used liability-threshold model [72]. The use of random effects in GeRSI is most beneficial for diseases that are known to be highly polygenic [72]. Lassosum is an example that estimates SNP effects using summary statistics and LD information from a reference panel in a penalized regression framework. Despite the need for parameter tuning, it has been shown to be computationally efficient and more accurate than the "clumping/pruning and thresholding" approach and LDpred [73].

LMM and Bayesian method make very different assumptions about the distribution of genetic effects and are expected to perform well in different situations. However, for a given disease, we typically do not know which assumptions will be more accurate. Motivated by this, hybrid methods such as SBayesR and PRS-CS were developed that combine the LMM and Bayesian method [74–77]. SBayesR performs Bayesian posterior inference of SNP weights through the combination of a likelihood that connects the multiple regression coefficients with summary statistics from GWAS and a finite mixture of normal distributions prior on the marker effects [75]. PRS-CS utilizes a high-dimensional Bayesian regression framework to infer

posterior effect sizes of SNPs and places the continuous shrinkage (CS) priors on SNP effect sizes, which can accommodate varying genetic architectures, provide substantial computational advantages, and enables multivariate modeling of local LD patterns [74]. When applied to predict common complex diseases, both of SBayesR and PRS-CS achieved improvement in prediction accuracy over LDpred and "clumping/pruning and thresholding" approach [74, 75].

In summary, development of optimal PRSs based on given GWAS data requires careful consideration of the threshold for SNP selection, weight assignment for selected SNPs, LD, and external knowledge (e.g., functional annotation and pleiotropic information). Practically, the optimal threshold, parameter, and statistical method that will be used for the construction of PRS can be determined based on the performance of the PRS in the study sample, or by using cross-validation techniques.

## Development of polygenic risk prediction models

Once the optimal PRS has been built, the next step is to develop a risk prediction model that incorporates the joint effects of the PRS and other established risk factors for a given cancer, including (but not limited to) lifestyle factors, environmental exposures, clinical characteristics, family history, and pathogenic mutations. Since established risk factors have substantial impact on cancer risk, and, more importantly, some of them can be potentially modifiable through lifestyle modifications or preventive interventions, the clinical utility of PRS should be fully evaluated along with non-genetic risk factors.

Detailed methodologies and considerations for the development of risk prediction models have been fully described in an in-depth review by Chatterjee N. et al. [3]. Briefly, common types of models include logistic regression model for case-control study and Cox proportional hazards model for cohort study. The development of models for the joint effect of PRS and other established risk factors requires characterization of the risk associated with individual risk factor, exploration of possible interactions between these factors, and assessment of the goodness-of-fit of the selected models. When a model has been built, it is important to produce absolute risk of developing the cancer for each individual [3].

Development of models for absolute risk involves the incorporation of data from various sources, including (but not limited to) epidemiological case-control and cohort studies, population-based disease and death registries, and national health surveys [3]. Fortunately, the establishment of large-scale cohorts and biobanks with deep phenotypes, such as the China Kadoorie Biobank (CKB), UK Biobank, All of Us Research Program, Biobank Japan, and Nordic efforts (e.g., in Danish, Estonian, Finnish, and other integrated biobanks), has largely facilitated the development of polygenic risk prediction models.

After the model development procedure to select the best risk prediction model, this model need to be validated in a representative sample (the validation sample) that is independent of the study population contributing to the model development procedure (the training sample) (Figure 1). Results from risk model evaluation should be reported for both the training sample and the validation sample for comparison [78, 79]. Ideally, prospective cohort studies are needed. Nested case-control studies, in which the cases emerge from a well-defined cohort and the controls are sampled from that same population, can also be used for the development and validation of risk prediction model [3, 80].

## Evaluation of polygenic risk prediction models

The utility of a risk prediction model is often evaluated by determining whether this model, which usually incorporates the joint effects of PRS and other risk factors for a disease, accurately stratifies the population into categories with sufficient distinct degrees of absolute risk to guide clinical or personal decision-making, such as preventive interventions or disease screening for populations at high risk. For binary diseases, it is conventional to report the distribution of the PRS, as well as the proportion of variance explained ($R^2$) and effect size estimates (e.g., odds ratios or hazards ratios from regression models) of the PRS and/or risk models [78, 79]. Subsequent steps for evaluation of a risk prediction model include judging how well the model differentiates those who will develop the disease in the future from those who will not (discrimination), and how similar the predicted risk is to the observed risk for individuals in different risk strata or with different risk factor profiles in the underlying population (calibration) [81]. Then the utility of a risk prediction model should further be evaluated for its risk stratification capacity, such as the proportions in which the population is stratified into

clinically relevant risk categories [82]. It is also important to keep models up to date when new information about risk factors and incidence rates of disease becomes available [3]. The added value of new risk factors incorporated into a model has been increasingly common to be assessed by net reclassification index (NRI) [81]. Recommended steps for evaluating the performance of polygenic risk prediction models are summarized in Table 1.

## Discrimination

Discrimination is a measure of how well the model can separate those who will develop the disease in the future and those who will not, which is of most interest when the primary purpose is the separation of diseased from non-diseased individuals, such as in diagnostic or screening testing [83]. The discrimination ability of a risk prediction model is commonly quantified by calculating the area under the receiver operating characteristic (ROC) curve or the concordance-statistic (c-statistic), which is defined as the probability that the predicted risk is higher for a randomly selected individual with a disease than a randomly selected individual without the disease [84]. The area under the receiver operating characteristic curve (AUC), or c-statistic, ranges from 50% to 100%. The greater the AUC, the better the risk discriminatory ability of the model. An AUC of 100% corresponds to perfect discrimination, which will be achieved if the predicted risks or scores for patients with a disease are always higher compared with those without the disease [3, 81, 83, 85]. There are several scales for the interpretation of AUC or c-statistic. A generally accepted scale suggests that an AUC or c-statistic of less than 0.60 means poor discrimination; 0.60–0.75 means possibly helpful discrimination; and above 0.75 means clearly useful discrimination [81].

Although discrimination is an important characteristic in the evaluation of model performance, sole reliance on the ROC curves would seem inappropriate. ROC curves provides no information regarding the accuracy of the absolute risks that the model predicts [82]. Sometimes, a risk prediction model can have a high AUC or c-statistic, but still provide misleading absolute risks [86]. In addition, the ROC curve and the c-statistic can be insensitive to important changes in absolute risk estimates [87]. For instance, the addition of a new risk factor may contribute important prediction power to a model that is clinically meaningful. However, the AUC may not increase substantially, particularly in circumstances where the initial AUC is already high [87]. Therefore, discrimination ability

should be used in combination with other measures to evaluate the performance of risk prediction models.

## Calibration

Calibration is often considered as the most important property of a risk prediction model, particularly in those used to estimate population disease burden and to plan population-based interventions. Calibration reflects the ability of a model to correctly estimate the risk for subjects with different risk factor profiles in the underlying population (i.e., if model-predicted risk of a disease agrees with the observed risk) [81, 82]. Subjects can be classified into strata based on their predicted risks or combinations of predictors. Then calibration can be assessed by comparing predicted and observed risk within these different strata [88]. Graphical representation of the relationship between predicted and observed risk is the best way to evaluate model calibration [81]. An alternative is to use statistical tests to determine whether the difference between the predicted and the observed risk of disease can be explained by chance. The Hosmer–Lemeshow test is a commonly used statistical test of model calibration. A statistically significant result (e.g., p-value <0.05) suggests that the difference between observed and predicted risks cannot be explained by chance, implying poor calibration [82]. However, there are limitations with the Hosmer–Lemeshow test. A statistically significant result provides no indication of the magnitude of the difference or whether it varies among individuals with low vs. high risk. The Hosmer–Lemeshow test may also be confounded by sample size. When the sample size is large, a clinically trivial difference could lead to a statistically significant result [81].

As risk prediction is likely to be most clinically relevant for individuals at extremely high or low risk, evaluating the adequacy of models at extreme level of risk requires particular attention for clinical application [89]. Sometimes, calibration can be good in some individuals, but not as good in others. For instance, a model can be accurate at estimating risk for individuals in the 0–20% of the risk score distribution, but overestimates risks in individuals with higher risk score. Consequences of such poor calibration among individuals at higher risk depend on the thresholds of clinically relevant risk categories: if the threshold for clinical intervention is below 20%, for example, this model would still have clinically utility because overestimation among individuals with risk greater than 20% would be irrelevant [81, 90].

## Risk stratification

Having established that a model is well-calibrated in the underlying population, it needs to be further assessed for its utility in clinical or public health applications. As noted earlier, the utility of a risk prediction model generally depends on its risk stratification ability to divide the population into distinct categories of absolute risk on which clinical or personal decision-making can be based [3, 4, 82]. The risk stratification ability of a model depends on how much variation in estimated risk it can provide in an underlying population. Consider, for example, a model that includes age as the only predictor for lung cancer risk. The risk of a person at a given age may be estimated by the average risk of lung cancer for the whole population at his age, potentially using data from population-based registries. As more risk factors are identified (e.g., smoking and genetic risk factors) and included in the model, estimated risks will be more variable between individuals and a larger proportion of individuals could be stratified into more extreme risk categories [3, 82], so that individuals at high risk can be offered targeted screening and interventions to address their risks of developing the cancer, and individuals at low risk can avoid unnecessary screening and interventions.

The risk stratification capacity of a model can be quantified by the proportions of individuals who are allocated into clinically relevant risk categories. For screening or other interventions targeted to the high risk population, we may also evaluate the proportion of patients who will develop a disease in the future that may be identified as being at high risk [91, 92]. An ideal model would be capable of identifying a small fraction of the whole population that will give rise to the majority of future diseases. However, as a substantial proportion of future diseases can still arise outside the groups identified as exceeding a certain risk threshold, the more realistic goal is to identify a subset of individuals at elevated risk of disease [93]. The absolute risk thresholds to determine how individuals should be assigned to distinct risk categories will depend on the benefits and harms of an intervention strategy in the underlying population [3]. The medical community has already defined the recommended risk threshold that justifies certain medical interventions for several cancers: the more risky the intervention, the higher the level of absolute risk threshold [4].

## Net reclassification index

The NRI is commonly used to quantify whether new risk factors provide clinically relevant improvements in risk prediction. Net reclassification involves classifying individuals into risk strata and quantifying the degree to which the new model can provide more accurate classification compared with a previous model, that is, shift individuals who have or will have an event to higher risk categories and individuals who do not have or will not have the event to lower risk categories [94, 95]. Several methods of calculating the NRI have been proposed. The widely recommended method is to calculate the category-based NRI, with the formulas shown in Table 1 [81, 96–98].

When clinical guidelines have recommended risk threshold for a given intervention, such as an initiation of chemoprevention or screening, it is strongly recommended that the risk categories be defined using clinically meaningful threshold. As these thresholds are supported by cost-effectiveness analyses, the NRI captures the change in a person's predicted risk that crosses one of such thresholds and thus translates into a meaningful change in clinical decisions [96].

The overall NRI implicitly weights for the event rate, with one/event rate and one/(one – event rate) serving as costs for false-negative results (persons with the event classified downward) and false-positive results (persons without the event classified upward), respectively [95, 99]. However, a different weighting of false-positive and false-negative results is often more clinically appropriate [100]. This can be incorporated in a weighted version of the NRI if the event NRI and nonevent NRI are presented separately or when a reclassification table is provided [94, 101]. If the misclassification of persons with the event leads to more serious consequences than the misclassification of those without the event, more weight might be assigned to the classification of individuals with the event, and less weight assigned to the classification of individuals without the event [102].

# Polygenic risk scores for cancer risk prediction

Development and evaluation of PRS is an area of active research with the potential to improve disease risk assessment and guide clinical decision-making about selecting appropriate interventions to manage disease risk. In this section, we discuss the research progress in the risk prediction of common cancers by using PRS, and illustrate several studies that performed cross-cancer evaluation of PRS. Since current PRSs are primarily developed and evaluated in European populations, unless otherwise stated, study participants are of European ancestries.

## Lung cancer

To date, GWASs have collectively identified 51 susceptibility loci for lung cancer [103, 104]. These GWASs were conducted in European and Asian populations, which revealed both shared and population specific genetic etiology of lung cancer. To evaluate the utility of PRS in identifying high risk populations of lung cancer for prevention, Dai et al. built a PRS specific for Chinese populations using a set of 19 SNPs ($PRS_{19}$) and evaluated its performance at predicting lung cancer incidence in an independent prospective cohort of 95,408 individuals randomly selected from CKB [103]. They reported that individuals in the top 10% of $PRS_{19}$ distribution had 1.96-fold increased risk for developing lung cancer compared with those in the lowest 10% [103]. When using the top 5%, 5–95%, and bottom 5% of PRS to define high, intermediate, and low genetic risk populations, consistently separate curves of lung cancer events were observed in the CKB cohort during follow-up, with a relative risk of lung cancer being 137% higher among participants at high genetic risk than those at low genetic risk [103].

Another study constructed the European ancestry-based PRS using 128 SNPs (PRS-128) and evaluated the PRS in an independent cohort from UK Biobank [105]. It was reported that, in the UK Biobank, individuals in the top 10% of the PRS distribution had 2.39-fold greater risk for lung cancer compared with those in the bottom 10%. When incorporating the PRS into the $PLCO_{all2014}$ model [106] for predicting the risk of lung cancer, an improvement in the discriminatory ability was observed, with the AUC increased from 0.828 to 0.832 [105]. An assessment of risk trajectory showed that PRS distribution can affect when the individuals reach the absolute risk threshold for the low-dose computed tomography (LDCT) screening. For instance, assuming a 1.5% five year risk threshold for the LDCT screening, individuals who smoked but without family history reach the threshold at age 61, whereas those who are at the top 1% of PRS distribution would reach the threshold at age 53. Among those who smoked and with positive family history of lung cancer, the average age to reach the 1.5% five year risk threshold would be 56, but those who are at top 5% of PRS distribution would reach the threshold at age 52 [105, 107].

Current studies for lung cancer demonstrated the ability of PRS in optimizing the definition of high risk populations beyond smoking and other known risk factors, which has potential utility in informing the optimal lung cancer LDCT screening strategy. Further refinement of the algorithm for PRS construction is needed as studies with larger sample sizes are continuing to discover novel susceptibility loci and other risk factors for lung cancer.

## Breast cancer

In the field of breast cancer, large-scale studies and reproducible methods have facilitated the development and validation of polygenic risk prediction models in European populations. Studies that examined the effect of PRS have consistently reported positive associations between PRS and breast cancer risk [16, 108–114]. Although several PRSs have been developed, one of the current best performing PRS incorporates information from 313 SNPs ($PRS_{313}$) [16]: compared with women in the middle quintile (40th–60th percentile at population average risk), those in the highest 1% of the $PRS_{313}$ distribution had approximately four-fold greater risk for breast cancer [16]. $PRS_{313}$ had moderate discriminatory ability (AUC=0.630, 95% confidence interval: 0.628–0.651) and was well-calibrated, which predicts breast cancer risk accurately in the tails of the distribution. The estimated lifetime absolute risk for breast cancer by age 80 years ranged from approximately 2.5% for women in the lowest centile of the $PRS_{313}$ to 32.6% for those in the highest centile [16]. Recent studies have used new statistical methods, such as LDpred and PRS-CS, for the development of PRS with the goal of improving risk prediction [109, 115]. Nevertheless, further research is needed to evaluate the performance of these PRSs and to determine the optimal PRS for breast cancer.

The risks conferred by PRS differ between breast cancer subtypes, with risk stratification of ER-positive disease being more effective than ER-negative disease [16, 116–120]. For example, the estimated lifetime absolute risk by age 80 years ranged from 2% for women in the lowest centile of the $PRS_{313}$ distribution to 31% for those in the highest centile for ER-positive disease, while the absolute risks ranged from 0.55% to 4% for ER-negative disease [16]. As findings from GWASs highlight heterogeneity across breast cancer subtypes [10, 121], it is necessary to build subtype-specific PRS to improve its prediction accuracy [16].

Several studies have shown a potential role of PRS in refining risk estimates for carriers of high-to-moderate penetrant mutations in breast cancer risk genes, with the PRS predicting substantial absolute risk differences for women at extremes of PRS distribution [122–125]. The effect of PRS on breast cancer risk is, to some extent, independent of family history. Although observed risk associated with PRS was attenuated in women with a family history of breast cancer, this association was observed for both

women with and without a family history [16, 114, 126–129]. PRSs have also been shown to be independent of other established risk factors for breast cancer, including mammographic density [111], lifestyle behaviors (diet, physical activity, smoking, alcohol consumption, BMI, waist circumference) [108, 117, 130], reproductive factors (age at menarche, parity, number of children, age at first full-term pregnancy, and breastfeeding) [108, 117], and exogenous hormonal factors (use of oral contraceptives and use of postmenopausal hormone replacement therapy) [108, 117]. These findings suggest that the PRSs are needed to be considered along with pathogenic mutations, family history, and other established risk factors to build an optimal risk prediction model for breast cancer.

Researchers have recently incorporated the PRS into existing risk prediction models for breast cancer, such as the Breast Cancer Risk Assessment Tool (BCRAT, also known as the Gail model) [119, 131–134], Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) [135, 136], International Breast Cancer Intervention Study model (IBIS, also known as the Tyrer–Cuzick model) [111, 116, 118, 119], BRCAPRO [119], and Breast Cancer Surveillance Consortium (BCSC) [113, 137]. Currently, all studies have reported improved discrimination for predicting breast cancer when PRS is added to the model. A study in high risk women found that adding a panel of 88 susceptibility SNPs ($PRS_{88}$) to the IBIS model resulted in a substantial improvement in the predictive power, with the c-statistic increased from 0.55 to 0.60 [118]. When reweighted to the original population, the percentage of women with 10 year predicted risk above 8% was 18% for IBIS model and increased to 21% if recalibrated $PRS_{88}$ was added [118]. BOADICEA uses information on family history, high-to-moderate pathogenic variants in breast cancer genes, tumor pathology, and basic demographic factors (such as age and ethnicity) to estimate the future risks of developing breast or ovarian cancer [138]. Lee et al. recently extended the BOADICEA model for breast cancer (BOADICEA v5 model) to incorporate the $PRS_{313}$ [16], mammographic density, and other risk factors [136]. They demonstrated that the $PRS_{313}$ provides greater level of risk stratification in the population than epidemiological risk factors alone, and that the greatest breast cancer risk stratification is achieved by using the combined effects of $PRS_{313}$ and epidemiological risk factors in the BOADICEA v5 model [136]. Lakeman et al. further validated this model in a Dutch population: the best discrimination was achieved again when $PRS_{313}$ and epidemiological risk factors were considered jointly, with the largest contribution deriving from the $PRS_{313}$ [135].

These findings highlight the potential of PRSs for improving discrimination ability of current breast cancer risk prediction models, although even the best models described to date leave room for continued improvement [139]. It should be noted that several studies reported calibration issues in polygenic risk prediction models [118, 135]. For example, when applied to a Dutch population, the BOADICEA v5 model for breast cancer underestimated the observed risks, especially in the higher categories of risk [118, 135]. Therefore, application to target population would require recalibration of the model for accurate risk assessment.

## Colorectal cancer

CRC is the third most common cancer and the second leading cause of cancer-related death worldwide [2], yet it is more suitable for screening and prevention than any other malignancy because CRC screening is effective for detecting cancer at an earlier stage and for reducing cancer risk by removing premalignant lesions [140]. PRSs that aggregate the increasing number of known genetic susceptibility variants identified by GWASs have been developed and evaluated for the prediction of CRC risk to inform screening and other prevention strategies [12, 57, 60, 141–147]. The discriminatory ability has been improved as more susceptibility variants are included. Currently, one of the best performing model is an LDpred-derived PRS including nearly 1.2 million genetic variants (AUC=0.654) [60]. Individuals in the top 1% of this LDpred-derived PRS distribution had 2.68-fold increased CRC risk compared with the remaining 99% of the individuals [60].

Associations between PRSs and CRC risk has been shown to be independent of lifestyle factors [148–151] and colonoscopy status [146, 147]. A combination of these factors allows for tailored recommendation for the starting age of screening [60, 141, 147]. For example, to inform the optimal age to begin screening, Jeon et al. developed a risk prediction model based on family history, 19 lifestyle and environmental factors (E-score), as well as 63 susceptibility SNPs of CRC (G-score) [141]. Risk assessment based on this model may help to personalize the age of the onset of screening: in individuals with no family history of CRC, the starting ages calculated based on combined E-score and G-score differed by 12 years for men and 14 for women, for individuals with the highest vs. the lowest 10% of risk [141]. In addition to informing the starting age of screening, PRS can also help to determine the length of screening interval after negative findings from colonoscopies [145]. For instance, Guo et al. evaluated CRC risk based on a 90-SNP

PRS and time since last negative colonoscopy. They utilized tertiles of PRSs among controls to classify individuals into low, medium, or high genetic risk. For individuals who had a negative finding from colonoscopy, a 42–85% lower risk of CRC was observed within 10 years compared to individuals without colonoscopy and with low genetic risk. Beyond 10 years after a negative finding from colonoscopy, significantly lower risk only persisted for individuals with low and medium genetic risk, but not for those with high genetic risk. The authors concluded that the recommended 10 year screening interval for colonoscopy may not need to be shortened among people with high genetic risk, but could potentially be prolonged for people with low and medium genetic risk [145]. Finally, calculation of detailed absolute risk of CRC, based on PRS and other risk factors, is needed to facilitate risk communication and to better inform the public about the potentials and limits of cancer prevention [146].

There is evidence that the PRS risk gradient for CRC has a negative dependency on age. For instance, Archambault and colleagues [152] reported that a 95-SNP PRS was more strongly associated with early-onset than late-onset CRC. The OR per standard deviation went from 1.75 to 1.44 as age increased from the 40s to 70s and that this trend was highly significant (p=$3.44 \times 10^{-10}$) [152]. On the log (OR) scale, these risk gradients are from 0.56 to 0.37, which is an almost linear decline of approximately 0.06 per decade [152–154]. These observations suggest that the PRS, along with other risk factors, might identify younger individuals who would benefit from tailored prevention strategies. On the other hand, however, using a PRS with an age-dependent risk gradient would lead to underestimates of the risk at younger ages and overestimates at older ages, which could compromise the calibration and therefore the value of the PRS for guiding screening [152–154].

## Gastric cancer

Gastric cancer remains one of the most common cancers worldwide, ranking fifth for incidence and fourth for mortality globally [2]. Incidence rates of gastric cancer are highest in Eastern Asia; more than 40% of new cases and deaths of gastric cancer occur in China [155]. Jin et al., developed a PRS derived from 112 SNPs for gastric cancer and evaluated its utility and effectiveness in an independent cohort of Chinese individuals (i.e., the CKB cohort) [156]. When using the top 20%, 20–80%, and bottom 20% of PRS to define high, intermediate, and low genetic risk categories, the PRS could identify subjects who are at a high risk of incident gastric cancer independent of

lifestyle factors in the CKB cohort: individuals with a high genetic risk had a two-fold increased risk of gastric cancer compared to those with a low genetic risk. Among individuals with a high genetic risk, those with a favorable lifestyle had a 47% lower risk of gastric cancer compared with those with an unfavorable lifestyle, with the 10 year absolute risk reduced from 1.62% to 0.49%. The 112 SNP-based PRS appears to be a practical and reliable genetic predictor for stratifying the risk of gastric cancer in Chinese populations. In addition, an increased genetic risk can be offset by adhering to a healthy lifestyle [156].

## Cross-cancer evaluation of PRS

Systematically assessing the added value of genetic information and examining how it affects lifetime risk trajectories are important to realize the promise of PRS in cancer risk assessment [157–160]. A recent study quantified the added value of integrating cancer-specific PRS with family history and modifiable risk factors for 16 cancers [160]. Incorporating PRS improves discrimination ability for all 16 cancers examined, but the magnitude of this improvement varies substantially. The largest increase in the c-index was observed for testicular, thyroid, prostate, breast, and CRCs. However, modeling the PRS in addition to conventional risk factors yielded marginal improvements in discrimination for cancers of the lung, endometrium, bladder, oral cavity/pharynx, and kidney. These cancers have strong environmental risk factors, such as smoking, alcohol consumption, obesity, and human papillomavirus (HPV) infection. The improvement in discrimination ability also translated into a refinement of risk stratification after accounting for conventional risk factors, as illustrated by more divergent five year risk trajectories. For certain cancers, such as melanoma, breast, colorectal, and pancreatic cancers, PRS was the primary determinant of risk stratification. For others, such as lung and bladder cancers, modifiable risk factors had a stronger impact on five year risk trajectories [160].

Another important question remains about how far we can improve the predictive performance using genome-wide genetic data. To this end, Zhang et al. showed that the theoretical maximal AUC with the best achievable PRS, based on genome-wide data of common variants, varies from 64% (endometrial and ovarian cancer) to 88% (testicular cancer) and in the range of 70–80% for most cancers [161]. The theoretical maximal relative risk for subjects at the 99th risk percentile of the best achievable PRS, compared to average risk, ranges from 12 for testicular to 2.5 for ovarian cancer. Across cancer types, PRSs show

varying levels of risk stratification. For common cancers, such as breast, colorectal, and prostate, a PRS with even modest discriminatory power can provide substantial risk stratification in the population. In contrast, for testicular cancer, even though its PRS could yield a higher AUC (e.g., in the range 80–90%), the degree of risk stratification will be modest because of the low incidence of this cancer [161].

Taken together, incorporating PRS with established risk factors can improve risk prediction for common cancers. The prediction power could be further improved by refining prediction models, using larger number of genetic variants, or larger sample sizes. Current GWAS-based PRS does not include rare variants. As DNA sequencing technologies will facilitate the discovery of cancer-associated rare variants [162], it can be expected that, in the future, incorporation of rare variants can further improve the predictive performance of polygenic prediction models.

# The utility of polygenic risk scores in cancer screening and precision prevention

Future application of PRS in public health and clinical practice holds significant promise. In this section, we will illustrate the utility of PRS in cancer prevention, including quantitating disease risk in various subpopulations to take strategies to prevent or delay the onset of disease (primary prevention), and identifying high risk individuals who are eligible for cancer screening (secondary prevention). The clinical utility of PRS depends not only on its risk stratification ability but also on the availability of appropriate risk-reducing interventions as well as the complex interplay between disease-specific and intervention-specific risks and benefits [3, 4, 8].

## Cancer screening

PRS has the potential to identify a larger fraction of the population who might benefit from the population-based screening programs for the early detection of cancer. Most recently, one study of CRC evaluated the risk stratification ability of an LDpred-derived PRS including nearly 1.2 million genetic variants [60]. This PRS was able to identify the top 30% of the study population as having risk for CRC similar to those with an affected first-degree relative, for whom some guidelines recommend initiation of screening with colonoscopy at an earlier age. It should be noted that 89.5% of these individuals who were in the top

30% risk based on the LDpred-derived PRS have no family history of CRC and would have been classified as average risk under current screening guidelines, but might benefit from earlier screening [60].

Findings have also demonstrated the utility of PRS in tailored recommendation of cancer screening, rather than simply defining a high risk group that is itself heterogeneous. A recent study evaluated the impact of $PRS_{313}$ on the starting age of screening for breast cancer [16]. For instance, women in the United Kingdom will become eligible to enter the mammographic screening program when they turn 47 years old. The average 10 year absolute risk of breast cancer for woman at this age is 2.6% in the general population. The study found that women in the top 1% of genetic risk, according to the $PRS_{313}$, would reach this risk threshold in their early 30s, whereas women in the bottom 20% of the $PRS_{313}$ distribution would remain below this threshold up until age 80 years [16]. Similar results have been shown in CRC [141], in which a 63-SNP PRS, in conjunction with lifestyle and environmental factors, can have a substantial impact on the starting age of colonoscopy screening. For individuals with no family history, the recommended age to initiate screening is approximately 12 years earlier for individuals at the top 10% of the risk score (44 for men and 50 for women) than those at the bottom 10% of the risk score (56 for men and 64 for women) [141].

## Precision prevention

Individualized management of disease is essential for precision medicine, with genetic information often used to facilitate personalized healthcare [163]. The potential utility of PRSs in prioritizing chemoprevention is illustrated by a recent study relating breast cancer PRSs to the use of risk-lowering therapies [164]. Randomized controlled trials (RCTs) evaluating anti-estrogens in primary prevention of breast cancer have consistently reported a reduced incidence in hormone receptor-positive subtypes of the disease [165]. Accordingly, the US Preventive Services Task Force recommend the initiation of risk-reducing medications (such as tamoxifen, raloxifene, or aromatase inhibitors) in asymptomatic women aged ≥35 years at increased risk for breast cancer and low risk for adverse medication effects [166]. Although there is no single cutoff for defining increased risk for all women, those with a five year absolute risk for breast cancer above 3% are likely to obtain more benefit than harm from risk-reducing medications and should be offered preventive medications if their risk of harms is low [166, 167]. Hurson et al. performed

risk prediction for breast cancer using a combination of the $PRS_{313}$ and classical risk factors. They showed that addition of the $PRS_{313}$ resulted in the reclassification of 9.2% of US women moving from below the 3% five year risk threshold to above, translating into a stronger recommendation for risk-reducing medications [165]. Therefore PRSs are useful, independent of classical risk factors, for the identification of individuals at elevated risk who could receive greater benefit from targeted risk-reducing strategies under current clinical guidelines.

Another area where PRS might be useful is for communicating benefits to individuals regarding targeted lifestyle interventions. Avoiding certain risk factors (e.g., hormone replacement therapy for breast cancer), as well as adopting healthier lifestyle patterns (e.g., smoking cessation, alcohol intake reduction, exercise, and maintaining a healthy weight), can have long-term cancer-preventive effects [2]. In support of this, Kachuri et al. quantified the predictive value of integrating cancer-specific PRS with family history and modifiable risk factors for 16 cancers, which indicated that individuals at highest levels of genetic risk may also experience larger decreases in risk from shifting to a healthier lifestyle [160]. Dai et al. evaluated a PRS composed of 19 GWAS-identified risk SNPs of lung cancer in the CKB cohort and showed that heavy smokers within the highest 5% of genetic risk can offset much of their risk by not smoking throughout their lifetime, leading to a reduction of their lung cancer risk by 62% in the study population [103]. For breast cancer, if healthy lifestyle choices were employed, those in the top tertile of the $PRS_{313}$ distribution would have 27% and 32% reductions in their risk of invasive breast cancer for premenopausal and postmenopausal women, respectively [130]. Promisingly, when genetic risk of complex disease is returned to high risk individuals, potentially positive behavioral changes have been observed [168]. However, we still lack experience of how to use PRS to motivate behavior change [169].

# Challenges and future perspectives

## PRS and environment interaction

Nearly all the common cancers have both genetic and environmental risk factors. Identifying polygenic interactions with environmental factors becomes an attractive research area. The existence of PRS-environment interactions implies different effect of the PRS on disease risk in individuals with different environmental exposures [170]. With a positive interaction, the effect of high PRS would be amplified in the presence of an environmental risk factor (or a combination of

environmental risk factors), putting this subgroup of the population at particularly high risk of disease [169]. These individuals, identified by PRS and environmental factors, could form a specific target group for cancer screening or other prevention strategies [169]. Modeling and testing PRS-environment interactions are essential for an accurate estimate of disease risk based on PRS, the environment, and their joint effect [171].

Numerous studies have assessed the interaction between PRS and environmental factors, providing mixed evidence of their interactive effect [16, 108, 117, 130, 149–151]. For instance, one study of breast cancer tested the null hypothesis of multiplicative joint associations for a 77-SNP PRS and environmental risk factors, reporting evidence of PRS-environment interactions for alcohol consumption, adult height and use of menopausal hormone therapy in ER-positive breast cancer [117]. However, more recently, a study evaluating the $PRS_{313}$ for breast cancer reported null interactions [108].

Most previous studies used the multiplicative model to evaluate PRS-environment interactions, with additive interactions being rarely assessed. Additive interaction could provide more intuitive information for disease prevention, because it can help to identify groups of individuals who are more likely to benefit from interventions by directly quantifying the absolute risk reduction [169, 172]. Exploratory studies have reported additive interaction between polygenic risk and lifestyle factors in breast and CRC [130, 149, 173]. It should be noted that testing for PRS-environment interactions may not reach statistical significance because of insufficient power. Although model misspecification owing to the omission of interactions is unlikely to have a major impact on discriminatory ability, it can affect the calibration of models [174]. Therefore, goodness-of-fit tests should be performed when assessing the adequacy of models.

## Applicability of PRS across ethnic groups

One of the major challenges surrounding clinical implementation of PRS is to ensure that they are equally applicable to individuals across ethnic groups to avoid exacerbating health disparities [175]. Since the majority of GWAS and imputation reference panels have currently been in European populations, PRSs are primarily developed and evaluated in individuals of European descent, which usually leads to a decrease in predictive accuracy when applied to non-European ancestries [176, 177]. This lack of transferability is thought to be attributed to various reasons: (a) differences in SNP allele frequencies (GWAS

favor the discovery of genetic variants that are common in the study population) and LD patterns between ancestries; (b) confounding due to population stratification in the GWAS; and (c) differences in the true genetic architectures of disease, including gene-environment interactions [175–177].

Studies have begun to assess whether PRS generated from one ethnic group may be predictive of the same disease in another. Indeed, for breast cancer, studies assessing European ancestry-based models in women of Hispanic, African American, and African ancestries found that they generally had lower discriminatory ability for breast cancer risk prediction than that reported in European populations [119, 134, 178]. On the other hand, the European ancestry PRSs have been found to have similar performance in Latinas [62] and Asians [179]. As such, more systematic and thorough evaluations for the utility of PRS in clinical settings across multiple ancestries are still needed. We must be cautious in the implementation of PRSs so as not to provide inaccurate information and exacerbate existing health disparities.

To avoid exacerbating health disparities, promoting large-scale GWAS in homogenous populations from understudied ethnic groups to generate ancestry-specific PRS will be a potential path forward [180]. Substantial investment will be needed to acquire sufficiently large sample sizes for PRS to achieve equal performance in other ethnic groups [47, 175]. Another path toward parity in PRS accuracy is increasing diversity among study participants included in genetic research [175]. Key methodological considerations for GWAS in ethnically diverse populations have been recently discussed, including the choice between performing association analysis in single ancestry groups followed by meta-analysis and performing mixed model for multi-ancestry groups [181]. Accordingly, novel computational methods that bring in data from ethnically diverse populations to build trans-ancestry PRS, such as MultiPred, are being developed [182–185]. Despite the difficulties in analyzing genetic data in ethnically diverse populations, it is scientifically and clinically imperative, and there are a growing number of analytical methodologies to do it well [181].

## Validation of PRS for its clinical utility

Standardization of any risk stratification tool is essential for consistent implementation across settings. Currently, standards and methods to build PRS are constantly developing. The prediction performance of PRSs for a disease can vary depending on the number of SNPs included,

the SNP weight estimates based on a certain GWAS dataset utilized, the specific computational method used for PRS construction and handling of LD, and the training dataset used to determine the optimal PRS. Prediction performance can also vary owing to the covariates, such as age and sex, adjusted for in the assessment of PRSs. This inconsistency in the development and evaluation of PRS now becomes a major challenge during its clinical application. Therefore analysis guidelines and reporting standards, as well as additional resources, such as accessible database of published PRSs and external validation data sets, are necessary to improve comparability and evaluation of PRSs [8, 79, 186].

The dramatic decrease in the costs of genetic testing has increased the feasibility of applying PRSs in many clinical settings. The clinical utility of PRSs in personalized prevention, however, needs further validation in large-scale population-based prospective cohort studies. Ultimately, the impact of PRS on clinical decision-making and outcomes (e.g., reduction in incidence and mortality for high risk population and avoidance of over-diagnosis) must be carefully evaluated in pragmatic clinical trials (PCTs) and, when possible, RCTs prior to implementation in clinical settings [187, 188]. Finally, implementation of PRS into population screening programs also requires consideration of the social, ethical, and psychological outcomes. For example, consideration needs to be given to the acceptance and adoption of new risk-stratification programs that use genetic information (particularly for those with a reduced risk), training health professionals in developing best risk communication tools, and cost-effectiveness and cost-benefit evaluations of alternative prevention strategies [114, 187, 189].

## Concluding remarks

The past two decades have witnessed the great success of GWAS that are revolutionizing our understanding of cancer genetics. Taking advantage of the findings from GWAS coupled with more assessable and affordable array-based genotyping technologies, it is right time to develop polygenic risk prediction models and to evaluate its utility in the population level. However, future studies will be required to further refine PRS methodologies, better integrate both genetic and conventional risk factors of cancer, and to address other pending challenges surrounding the implementation of PRS to fully realize the promise of precision prevention of cancer.

# References

1. Global Health Estimates 2020. Deaths by cause, age, sex, by country and by region, 2000-2019. Geneva: World Health Organization; 2020.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA A Cancer J Clin 2021;71:209–49.
3. Chatterjee N, Shi J, Garcia-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. Nat Rev Genet 2016;17:392–406.
4. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. Nat Rev Genet 2018;19:581–90.
5. Britt KL, Cuzick J, Phillips KA. Key steps for effective breast cancer prevention. Nat Rev Cancer 2020;20:417–36.
6. Win AK, Macinnis RJ, Hopper JL, Jenkins MA. Risk prediction models for colorectal cancer: a review. Cancer Epidemiol Biomarkers Prev 2012;21:398–410.
7. Gray EP, Teare MD, Stevens J, Archer R. Risk prediction models for lung cancer: a systematic review. Clin Lung Cancer 2016;17: 95–106.
8. Lambert SA, Abraham G, Inouye M. Towards clinical utility of polygenic risk scores. Hum Mol Genet 2019;28:R133–R42.
9. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 2019;47:D1005–12.
10. Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, et al. Association analysis identifies 65 new breast cancer risk loci. Nature 2017;551:92–4.
11. Zhang H, Ahearn TU, Lecarpentier J, Barnes D, Beesley J, Qi G, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. Nat Genet 2020;52:572–81.
12. Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, et al. Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet 2019;51:76–87.
13. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic

14. susceptibility across histological subtypes. Nat Genet 2017;49: 1126–32.
14. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nat Genet 2018;50:928–36.
15. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 2018;50:1219–24.
16. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. Am J Hum Genet 2019;104:21–34.
17. Frank SA. Genetic predisposition to cancer - insights from population genetics. Nat Rev Genet 2004;5:764–72.
18. Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial risk and heritability of cancer among twins in Nordic countries. JAMA 2016;315:68–76.
19. Ahlbom A, Lichtenstein P, Malmstrom H, Feychting M, Hemminki K, Pedersen NL. Cancer in twins: genetic and nongenetic familial risk factors. J Natl Cancer Inst 1997;89:287–93.
20. Houlston RS, Peto J. Genetics and the common cancers. In: Eeles RA, Ponder BAJ, Easton DF, Horwich A, editors. Genetic predisposition to cancer. London: Chapman & Hall; 1996:208–26 pp.
21. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. Nature 1995;378:789–92.
22. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 1994;266:66–71.
23. Groden J, Thliveris A, Samowitz W, Carlson M, Gelbert L, Albertsen H, et al. Identification and characterization of the familial adenomatous polyposis coli gene. Cell 1991;66: 589–600.
24. Kinzler KW, Nilbert MC, Su LK, Vogelstein B, Bryan TM, Levy DB, et al. Identification of FAP locus genes from chromosome 5q21. Science 1991;253:661–5.
25. Nishisho I, Nakamura Y, Miyoshi Y, Miki Y, Ando H, Horii A, et al. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. Science 1991;253:665–9.
26. Joslyn G, Carlson M, Thliveris A, Albertsen H, Gelbert L, Samowitz W, et al. Identification of deletion mutations and three new genes at the familial polyposis locus. Cell 1991;66:601–13.
27. Lammi L, Arte S, Somer M, Jarvinen H, Lahermo P, Thesleff I, et al. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. Am J Hum Genet 2004;74: 1043–50.
28. Bell DW, Gore I, Okimoto RA, Godin-Heymann N, Sordella R, Mulloy R, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. Nat Genet 2005;37:1315–6.
29. Rahman N. Realizing the promise of cancer predisposition genes. Nature 2014;505:302–8.
30. Pharoah PD, Dunning AM, Ponder BA, Easton DF. Association studies for finding cancer-susceptibility genetic variants. Nat Rev Cancer 2004;4:850–60.
31. Peto J, Collins N, Barfoot R, Seal S, Warren W, Rahman N, et al. Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. J Natl Cancer Inst 1999;91:943–9.

32. Easton DF. How many more breast cancer predisposition genes are there? Breast Cancer Res 1999;1:14–7.

33. Chubb D, Broderick P, Dobbins SE, Frampton M, Kinnersley B, Penegar S, et al. Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. Nat Commun 2016;7:11883.

34. Anglian Breast Cancer Study Group. Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Br J Cancer 2000;83:1301–8.

35. Lubbe SJ, Webb EL, Chandler IP, Houlston RS. Implications of familial colorectal cancer risk profiles and microsatellite instability status. J Clin Oncol 2009;27:2238–44.

36. Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, Peto J, et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. Br J Cancer 2002; 86:76–83.

37. Chakravarti A. Population genetics–making sense out of sequence. Nat Genet 1999;21:56–60.

38. Cardon LR, Bell JI. Association study designs for complex diseases. Nat Rev Genet 2001;2:91–9.

39. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science 2001;291: 1304–51.

40. International HapMap Consortium. The international HapMap project. Nature 2003;426:789–96.

41. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. Nature 2015;526:68–74.

42. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001;409:928–33.

43. Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, et al. High-throughput genotyping by whole-genome resequencing. Genome Res 2009;19:1068–76.

44. Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. Nat Rev Cancer 2017;17:692–704.

45. Tam V, Patel N, Turcotte M, Bosse Y, Pare G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet 2019;20:467–84.

46. Marigorta UM, Rodriguez JA, Gibson G, Navarro A. Replicability and prediction: lessons and challenges from GWAS. Trends Genet 2018;34:504–17.

47. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, et al. Genetic analyses of diverse populations improves discovery for complex traits. Nature 2019;570:514–8.

48. Zaitlen N, Kraft P. Heritability in the genome-wide association era. Hum Genet 2012;131:1655–64.

49. Evans LM, Tahmasbi R, Vrieze SI, Abecasis GR, Das S, Gazal S, et al. Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. Nat Genet 2018;50:737–45.

50. Speed D, Cai N, Johnson MR, Nejentsev S, Balding DJ. Reevaluation of SNP heritability in complex human traits. Nat Genet 2017;49:986–92.

51. Speed D, Holmes J, Balding DJ. Evaluating and improving heritability models using summary statistics. Nat Genet 2020; 52:458–62.

52. Hou K, Burch KS, Majumdar A, Shi H, Mancuso N, Wu Y, et al. Accurate estimation of SNP-heritability from biobank-scale data irrespective of genetic architecture. Nat Genet 2019;51:1244–51.

53. Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. Concepts, estimation and interpretation of SNP-based heritability. Nat Genet 2017;49:1304–10.

54. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82.

55. Dai J, Shen W, Wen W, Chang J, Wang T, Chen H, et al. Estimation of heritability for nine common cancers using data from genome-wide association studies in Chinese population. Int J Cancer 2017;140:329–36.

56. Sampson JN, Wheeler WA, Yeager M, Panagiotou O, Wang Z, Berndt SI, et al. Analysis of heritability and shared heritability based on genome-wide association studies for thirteen cancer types. J Natl Cancer Inst 2015;107:djv279.

57. Frampton MJ, Law P, Litchfield K, Morris EJ, Kerr D, Turnbull C, et al. Implications of polygenic risk for personalised colorectal cancer screening. Ann Oncol 2016;27:429–34.

58. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. J R Soc Med 2011;104:510–20.

59. Chasioti D, Yan J, Nho K, Saykin AJ. Progress in polygenic composite scores in Alzheimer's and other complex diseases. Trends Genet 2019;35:371–82.

60. Thomas M, Sakoda LC, Hoffmeister M, Rosenthal EA, Lee JK, van Duijnhoven FJB, et al. Genome-wide modeling of polygenic risk score in colorectal cancer risk. Am J Hum Genet 2020;107: 432–44.

61. Hughes E, Tshiaba P, Gallagher S, Wagner S, Judkins T, Roa B, et al. Development and validation of a clinical polygenic risk score to predict breast cancer risk. JCO Precis Oncol 2020;4. https://doi.org/10.1200/PO.19.00360.

62. Shieh Y, Fejerman L, Lott PC, Marker K, Sawyer SD, Hu D, et al. A polygenic risk score for breast cancer in US Latinas and Latin American women. J Natl Cancer Inst 2020;112:590–8.

63. Wu J, Pfeiffer RM, Gail MH. Strategies for developing prediction models from genome-wide association studies. Genet Epidemiol 2013;37:768–77.

64. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81:559–75.

65. Euesden J, Lewis CM, O'Reilly PF. PRSice: polygenic risk score software. Bioinformatics 2015;31:1466–8.

66. Seibert TM, Fan CC, Wang Y, Zuber V, Karunamuni R, Parsons JK, et al. Polygenic hazard score to guide screening for aggressive prostate cancer: development and validation in large scale cohorts. BMJ 2018;360:j5757.

67. Karunamuni RA, Huynh-Le MP, Fan CC, Thompson W, Eeles RA, Kote-Jarai Z, et al. Additional SNPs improve risk stratification of a polygenic hazard score for prostate cancer. Prostate Cancer Prostatic Dis 2021;24:532–41.

68. Karunamuni RA, Huynh-Le MP, Fan CC, Thompson W, Eeles RA, Kote-Jarai Z, et al. African-specific improvement of a polygenic hazard score for age at diagnosis of prostate cancer. Int J Cancer 2021;148:99–105.

69. Huynh-Le MP, Fan CC, Karunamuni R, Thompson WK, Martinez ME, Eeles RA, et al. Polygenic hazard score is associated with

prostate cancer in multi-ethnic populations. Nat Commun 2021;
12:1236.

70. Vilhjalmsson BJ, Yang J, Finucane HK, Gusev A, Lindstrom S,
Ripke S, et al. Modeling linkage disequilibrium increases
accuracy of polygenic risk scores. Am J Hum Genet 2015;97:
576–92.

71. Hu Y, Lu Q, Powles R, Yao X, Yang C, Fang F, et al. Leveraging
functional annotations in genetic risk prediction for human
complex diseases. PLoS Comput Biol 2017;13:e1005589.

72. Golan D, Rosset S. Effective genetic-risk prediction using mixed
models. Am J Hum Genet 2014;95:383–93.

73. Mak TSH, Porsch RM, Choi SW, Zhou X, Sham PC. Polygenic
scores via penalized regression on summary statistics. Genet
Epidemiol 2017;41:469–80.

74. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via
Bayesian regression and continuous shrinkage priors. Nat
Commun 2019;10:1776.

75. Lloyd-Jones LR, Zeng J, Sidorenko J, Yengo L, Moser G, Kemper
KE, et al. Improved polygenic prediction by Bayesian multiple
regression on summary statistics. Nat Commun 2019;10:5086.

76. Moser G, Lee SH, Hayes BJ, Goddard ME, Wray NR, Visscher PM.
Simultaneous discovery, estimation and prediction analysis of
complex traits using a Bayesian mixture model. PLoS Genet
2015;11:e1004969.

77. Zhou X, Carbonetto P, Stephens M. Polygenic modeling with
Bayesian sparse linear mixed models. PLoS Genet 2013;9:
e1003264.

78. Babb de Villiers C, Kroese M, Moorthie S. Understanding
polygenic models, their development and the potential
application of polygenic scores in healthcare. J Med Genet 2020;
57:725–32.

79. Wand H, Lambert SA, Tamburro C, Iacocca MA, O'Sullivan JW,
Sillari C, et al. Improving reporting standards for polygenic
scores in risk prediction studies. Nature 2021;591:211–9.

80. Ganna A, Reilly M, de Faire U, Pedersen N, Magnusson P,
Ingelsson E. Risk prediction measures for case-cohort and
nested case-control designs: an application to cardiovascular
disease. Am J Epidemiol 2012;175:715–24.

81. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ,
et al. Discrimination and calibration of clinical prediction
models: users' guides to the medical literature. JAMA 2017;318:
1377–84.

82. Janes H, Pepe MS, Gu W. Assessing the value of risk predictions
by using risk stratification tables. Ann Intern Med 2008;149:
751–60.

83. Cook NR. Use and misuse of the receiver operating characteristic
curve in risk prediction. Circulation 2007;115:928–35.

84. Pencina MJ, D'Agostino RB Sr. Evaluating discrimination of risk
prediction models: the C statistic. JAMA 2015;314:1063–4.

85. Gerds TA, Cai T, Schumacher M. The performance of risk
prediction models. Biom J 2008;50:457–79.

86. Barili F, Pacini D, Rosato F, Roberto M, Battisti A, Grossi C, et al.
In-hospital mortality risk assessment in elective and non-
elective cardiac surgery: a comparison between EuroSCORE II
and age, creatinine, ejection fraction score. Eur J Cardio Thorac
Surg 2014;46:44–8.

87. Cook NR, Ridker PM. Advances in measuring the effect of
individual predictors of cardiovascular risk: the role of
reclassification measures. Ann Intern Med 2009;150:795–802.

88. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ,
Steyerberg EW. A calibration hierarchy for risk models was
defined: from utopia to empirical data. J Clin Epidemiol 2016;74:
167–76.

89. Song M, Kraft P, Joshi AD, Barrdahl M, Chatterjee N. Testing
calibration of risk models at extremes of disease risk.
Biostatistics 2015;16:143–54.

90. Sartipy U, Dahlstrom U, Edner M, Lund LH. Predicting survival in
heart failure: validation of the MAGGIC heart failure risk score in
51,043 patients from the Swedish heart failure registry. Eur J
Heart Fail 2014;16:173–9.

91. So HC, Kwan JS, Cherny SS, Sham PC. Risk prediction of complex
diseases from family history and known susceptibility loci, with
applications for cancer screening. Am J Hum Genet 2011;88:
548–65.

92. Pfeiffer RM, Gail MH. Two criteria for evaluating risk prediction
models. Biometrics 2011;67:1057–65.

93. Park JH, Gail MH, Greene MH, Chatterjee N. Potential usefulness
of single nucleotide polymorphisms to identify persons at high
cancer risk: an evaluation of seven common cancers. J Clin Oncol
2012;30:2157–62.

94. Pencina MJ, D'Agostino RB Sr., Steyerberg EW. Extensions of net
reclassification improvement calculations to measure
usefulness of new biomarkers. Stat Med 2011;30:11–21.

95. Pencina MJ, D'Agostino RB Sr., D'Agostino RB Jr., Vasan RS.
Evaluating the added predictive ability of a new marker: from
area under the ROC curve to reclassification and beyond. Stat
Med 2008;27:157–72.

96. Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg
EW. Net reclassification improvement: computation,
interpretation, and controversies: a literature review and
clinician's guide. Ann Intern Med 2014;160:122–31.

97. Sheth T, Chan M, Butler C, Chow B, Tandon V, Nagele P, et al.
Prognostic capabilities of coronary computed tomographic
angiography before non-cardiac surgery: prospective cohort
study. BMJ 2015;350:h1907.

98. Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net
reclassification indices for evaluating risk prediction
instruments: a critical review. Epidemiology 2014;25:114–21.

99. Vickers AJ, Elkin EB, Steyerberg E. Net reclassification
improvement and decision theory. Stat Med 2009;28:525–6.

100. Greenland S. The need for reorientation toward cost-effective
prediction: comments on 'Evaluating the added predictive
ability of a new marker: from area under the ROC curve to
reclassification and beyond' by M. J. Pencina et al, Statistics in
Medicine (DOI: 10.1002/sim.2929). Stat Med 2008;27:199–206.

101. Van Calster B, Vickers AJ, Pencina MJ, Baker SG, Timmerman D,
Steyerberg EW. Evaluation of markers and risk prediction
models: overview of relationships between NRI and decision-
analytic measures. Med Decis Making 2013;33:490–501.

102. Vickers AJ, Pepe M. Does the net reclassification improvement
help us evaluate models and markers? Ann Intern Med 2014;
160:136–7.

103. Dai J, Lv J, Zhu M, Wang Y, Qin N, Ma H, et al. Identification of risk
loci and a polygenic risk score for lung cancer: a large-scale
prospective cohort study in Chinese populations. Lancet Respir
Med 2019;7:881–91.

104. Bosse Y, Amos CI. A decade of GWAS results in lung cancer.
Cancer Epidemiol Biomarkers Prev 2018;27:363–79.

105. Hung RJ, Warkentin MT, Brhane Y, Chatterjee N, Christiani DC, Landi MT, et al. Assessing lung cancer absolute risk trajectory based on a polygenic risk model. Cancer Res 2021;81:1607–15.

106. Tammemagi MC, Church TR, Hocking WG, Silvestri GA, Kvale PA, Riley TL, et al. Evaluation of the lung cancer risks at which to screen ever- and never-smokers: screening rules applied to the PLCO and NLST cohorts. PLoS Med 2014;11:e1001764.

107. Pinsky PF, Gierada DS, Hocking W, Patz EF Jr., Kramer BS. National lung screening trial findings by age: medicare-eligible versus under-65 population. Ann Intern Med 2014;161:627–33.

108. Kapoor PM, Mavaddat N, Choudhury PP, Wilcox AN, Lindstrom S, Behrens S, et al. Combined associations of a polygenic risk score and classical risk factors with breast cancer risk. J Natl Cancer Inst 2021;113:329–37.

109. Mars N, Widen E, Kerminen S, Meretoja T, Pirinen M, Della Briotta Parolo P, et al. The role of polygenic risk and susceptibility genes in breast cancer over the course of life. Nat Commun 2020;11:6383.

110. Brentnall AR, van Veen EM, Harkness EF, Rafiq S, Byers H, Astley SM, et al. A case-control evaluation of 143 single nucleotide polymorphisms for breast cancer risk stratification with classical factors and mammographic density. Int J Cancer 2020; 146:2122–9.

111. van Veen EM, Brentnall AR, Byers H, Harkness EF, Astley SM, Sampson S, et al. Use of single-nucleotide polymorphisms and mammographic density plus classic risk factors for breast cancer risk prediction. JAMA Oncol 2018;4:476–82.

112. Vachon CM, Pankratz VS, Scott CG, Haeberle L, Ziv E, Jensen MR, et al. The contributions of breast density and common genetic variation to breast cancer risk. J Natl Cancer Inst 2015;107. https://doi.org/10.1093/jnci/dju397.

113. Shieh Y, Hu D, Ma L, Huntsman S, Gard CC, Leung JW, et al. Breast cancer risk prediction using a clinical risk model and polygenic risk score. Breast Cancer Res Treat 2016;159:513–25.

114. Mavaddat N, Pharoah PD, Michailidou K, Tyrer J, Brook MN, Bolla MK, et al. Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst 2015;107:djv036.

115. Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, et al. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. Nat Med 2020;26:549–57.

116. Evans DGR, Harkness EF, Brentnall AR, van Veen EM, Astley SM, Byers H, et al. Breast cancer pathology and stage are better predicted by risk stratification models that include mammographic density and common genetic variants. Breast Cancer Res Treat 2019;176:141–8.

117. Rudolph A, Song M, Brook MN, Milne RL, Mavaddat N, Michailidou K, et al. Joint associations of a polygenic risk score and environmental risk factors for breast cancer in the Breast Cancer Association Consortium. Int J Epidemiol 2018;47: 526–36.

118. Cuzick J, Brentnall AR, Segal C, Byers H, Reuter C, Detre S, et al. Impact of a panel of 88 single nucleotide polymorphisms on the risk of breast cancer in high-risk women: results from two randomized tamoxifen prevention trials. J Clin Oncol 2017;35: 743–50.

119. Dite GS, MacInnis RJ, Bickerstaffe A, Dowty JG, Allman R, Apicella C, et al. Breast cancer risk prediction using clinical models and 77 independent risk-associated SNPs for women

aged under 50 years: Australian Breast Cancer Family Registry. Cancer Epidemiol Biomarkers Prev 2016;25:359–65.

120. Holm J, Li J, Darabi H, Eklund M, Eriksson M, Humphreys K, et al. Associations of breast cancer risk prediction tools with tumor characteristics and metastasis. J Clin Oncol 2016;34:251–8.

121. Milne RL, Kuchenbaecker KB, Michailidou K, Beesley J, Kar S, Lindstrom S, et al. Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. Nat Genet 2017;49:1767–78.

122. Kuchenbaecker KB, McGuffog L, Barrowdale D, Lee A, Soucy P, Dennis J, et al. Evaluation of polygenic risk scores for breast and ovarian cancer risk prediction in BRCA1 and BRCA2 mutation carriers. J Natl Cancer Inst 2017;109:djw302.

123. Barnes DR, Rookus MA, McGuffog L, Leslie G, Mooij TM, Dennis J, et al. Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. Genet Med 2020;22:1653–66.

124. Borde J, Ernst C, Wappenschmidt B, Niederacher D, Weber-Lassalle K, Schmidt G, et al. Performance of breast cancer polygenic risk scores in 760 female CHEK2 germline mutation carriers. J Natl Cancer Inst 2021;113:893–9.

125. Gallagher S, Hughes E, Wagner S, Tshiaba P, Rosenthal E, Roa BB, et al. Association of a polygenic risk score with breast cancer among women carriers of high- and moderate-risk breast cancer genes. JAMA Netw Open 2020;3:e208501.

126. Du Z, Gao G, Adedokun B, Ahearn T, Lunetta KL, Zirpoli G, et al. Evaluating polygenic risk scores for breast cancer in women of African ancestry. J Natl Cancer Inst 2021;113:1168–76.

127. Li H, Feng B, Miron A, Chen X, Beesley J, Bimeh E, et al. Breast cancer risk prediction using a polygenic risk score in the familial setting: a prospective study from the Breast Cancer Family Registry and kConFab. Genet Med 2017;19:30–5.

128. Muranen TA, Mavaddat N, Khan S, Fagerholm R, Pelttari L, Lee A, et al. Polygenic risk score is associated with increased disease risk in 52 Finnish breast cancer families. Breast Cancer Res Treat 2016;158:463–9.

129. Evans DG, Brentnall A, Byers H, Harkness E, Stavrinos P, Howell A, et al. The impact of a panel of 18 SNPs on breast cancer risk in women attending a UK familial screening clinic: a case-control study. J Med Genet 2017;54:111–3.

130. Arthur RS, Wang T, Xue X, Kamensky V, Rohan TE. Genetic factors, adherence to healthy lifestyle behavior, and risk of invasive breast cancer among women in the UK Biobank. J Natl Cancer Inst 2020;112:893–901.

131. Zhang X, Rice M, Tworoger SS, Rosner BA, Eliassen AH, Tamimi RM, et al. Addition of a polygenic risk score, mammographic density, and endogenous hormones to existing breast cancer risk prediction models: a nested case-control study. PLoS Med 2018;15:e1002644.

132. Darabi H, Czene K, Zhao W, Liu J, Hall P, Humphreys K. Breast cancer risk prediction and individualised screening based on common genetic variation and breast density measurement. Breast Cancer Res 2012;14:R25.

133. Dite GS, Mahmoodi M, Bickerstaffe A, Hammet F, Macinnis RJ, Tsimiklis H, et al. Using SNP genotypes to improve the discrimination of a simple breast cancer risk prediction model. Breast Cancer Res Treat 2013;139:887–96.

134. Allman R, Dite GS, Hopper JL, Gordon O, Starlard-Davenport A, Chlebowski R, et al. SNPs and breast cancer risk prediction for

African American and Hispanic women. Breast Cancer Res Treat 2015;154:583–9.

135. Lakeman IMM, Rodriguez-Girondo M, Lee A, Ruiter R, Stricker BH, Wijnant SRA, et al. Validation of the BOADICEA model and a 313-variant polygenic risk score for breast cancer risk prediction in a Dutch prospective cohort. Genet Med 2020;22:1803–11.

136. Lee A, Mavaddat N, Wilcox AN, Cunningham AP, Carver T, Hartley S, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. Genet Med 2019;21:1708–18.

137. Shieh Y, Hu D, Ma L, Huntsman S, Gard CC, Leung JWT, et al. Joint relative risks for estrogen receptor-positive breast cancer from a clinical model, polygenic risk score, and sex hormones. Breast Cancer Res Treat 2017;166:603–12.

138. Antoniou AC, Cunningham AP, Peto J, Evans DG, Lalloo F, Narod SA, et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. Br J Cancer 2008; 98:1457–66.

139. Yanes T, Young MA, Meiser B, James PA. Clinical applications of polygenic breast cancer risk: a critical review and perspectives of an emerging field. Breast Cancer Res 2020;22:21.

140. Sandouk F, Al Jerf F, Al-Halabi MH. Precancerous lesions in colorectal cancer. Gastroenterol Res Pract 2013;2013:457901.

141. Jeon J, Du M, Schoen RE, Hoffmeister M, Newcomb PA, Berndt SI, et al. Determining risk of colorectal cancer and starting age of screening based on lifestyle, environmental, and genetic factors. Gastroenterology 2018;154:2152–64.

142. Schmit SL, Edlund CK, Schumacher FR, Gong J, Harrison TA, Huyghe JR, et al. Novel common genetic susceptibility loci for colorectal cancer. J Natl Cancer Inst 2019;111:146–57.

143. Saunders CL, Kilian B, Thompson DJ, McGeoch LJ, Griffin SJ, Antoniou AC, et al. External validation of risk prediction models incorporating common genetic variants for incident colorectal cancer using UK Biobank. Cancer Prev Res 2020;13:509–20.

144. Li X, Timofeeva M, Spiliopoulou A, McKeigue P, He Y, Zhang X, et al. Prediction of colorectal cancer risk based on profiling with common genetic variants. Int J Cancer 2020;147:3431–7.

145. Guo F, Weigl K, Carr PR, Heisser T, Jansen L, Knebel P, et al. Use of polygenic risk scores to select screening intervals after negative findings from colonoscopy. Clin Gastroenterol Hepatol 2020;18: 2742–51.

146. Carr PR, Weigl K, Edelmann D, Jansen L, Chang-Claude J, Brenner H, et al. Estimation of absolute risk of colorectal cancer based on healthy lifestyle, genetic risk, and colonoscopy status in a population-based study. Gastroenterology 2020;159:129–38.

147. Hsu L, Jeon J, Brenner H, Gruber SB, Schoen RE, Berndt SI, et al. A model to determine colorectal cancer risk using common genetic susceptibility loci. Gastroenterology 2015;148:1330–9.

148. Carr PR, Weigl K, Jansen L, Walter V, Erben V, Chang-Claude J, et al. Healthy lifestyle factors associated with lower risk of colorectal cancer irrespective of genetic risk. Gastroenterology 2018;155:1805–15.

149. Choi J, Jia G, Wen W, Shu XO, Zheng W. Healthy lifestyles, genetic modifiers, and colorectal cancer risk: a prospective cohort study in the UK Biobank. Am J Clin Nutr 2021;113:810–20.

150. Erben V, Carr PR, Guo F, Weigl K, Hoffmeister M, Brenner H. Individual and joint associations of genetic risk and healthy lifestyle score with colorectal neoplasms among participants of screening colonoscopy. Cancer Prev Res 2021;14:649–58.

151. Chen X, Jansen L, Guo F, Hoffmeister M, Chang-Claude J, Brenner H. Smoking, genetic predisposition, and colorectal cancer risk. Clin Transl Gastroenterol 2021;12:e00317.

152. Archambault AN, Su YR, Jeon J, Thomas M, Lin Y, Conti DV, et al. Cumulative burden of colorectal cancer-associated genetic variants is more strongly associated with early-onset vs late-onset cancer. Gastroenterology 2020;158:1274–86.

153. Li S, Hopper JL. Age dependency of the polygenic risk score for colorectal cancer. Am J Hum Genet 2021;108:525–6.

154. Li S. Negative age-dependence of the polygenic risk score gradient for colorectal cancer. Gastroenterology 2021;160: 2214–5.

155. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA A Cancer J Clin 2016;66: 115–32.

156. Jin G, Lv J, Yang M, Wang M, Zhu M, Wang T, et al. Genetic risk, incident gastric cancer, and healthy lifestyle: a meta-analysis of genome-wide association studies and prospective cohort study. Lancet Oncol 2020;21:1378–86.

157. Choi J, Jia G, Wen W, Long J, Zheng W. Evaluating polygenic risk scores in assessing risk of nine solid and hematologic cancers in European descendants. Int J Cancer 2020;147:3416–23.

158. Fritsche LG, Patil S, Beesley LJ, VandeHaar P, Salvatore M, Ma Y, et al. Cancer PRSweb: an online repository with Polygenic Risk Scores for major cancer traits and their evaluation in two independent biobanks. Am J Hum Genet 2020;107:815–36.

159. Jia G, Lu Y, Wen W, Long J, Liu Y, Tao R, et al. Evaluating the utility of polygenic risk scores in identifying high-risk individuals for eight common cancers. JNCI Cancer Spectr 2020;4:pkaa021.

160. Kachuri L, Graff RE, Smith-Byrne K, Meyers TJ, Rashkin SR, Ziv E, et al. Pan-cancer analysis demonstrates that integrating polygenic risk scores with modifiable risk factors improves risk prediction. Nat Commun 2020;11:6084.

161. Zhang YD, Hurson AN, Zhang H, Choudhury PP, Easton DF, Milne RL, et al. Assessment of polygenic architecture and risk prediction based on common variants across fourteen cancers. Nat Commun 2020;11:3353.

162. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet 2014;95:5–23.

163. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015;372:793–5.

164. Hurson AN, Pal Choudhury P, Gao C, Husing A, Eriksson M, Shi M, et al. Prospective evaluation of a breast-cancer risk model integrating classical risk factors and polygenic risk in 15 cohorts from six countries. Int J Epidemiol 2021:dyab036. https://doi.org/10.1093/ije/dyab036.

165. Pashayan N, Antoniou AC, Ivanus U, Esserman LJ, Easton DF, French D, et al. Personalized early detection and prevention of breast cancer: ENVISION consensus statement. Nat Rev Clin Oncol 2020;17:687–705.

166. Owens DK, Davidson KW, Krist AH, Barry MJ, Cabana M, Caughey AB, et al. Medication use to reduce risk of breast cancer: US Preventive Services Task Force recommendation statement. JAMA 2019;322:857–67.

167. Freedman AN, Yu B, Gail MH, Costantino JP, Graubard BI, Vogel VG, et al. Benefit/risk assessment for breast cancer chemoprevention with raloxifene or tamoxifen for women age 50 years or older. J Clin Oncol 2011;29:2327–33.

168. Frieser MJ, Wilson S, Vrieze S. Behavioral impact of return of genetic test results for complex disease: systematic review and meta-analysis. Health Psychol 2018;37:1134–44.

169. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. Genome Med 2020;12:44.

170. Ottman R. Gene-environment interaction: definitions and study designs. Prev Med 1996;25:764–70.

171. Hunter DJ. Gene-environment interactions in human diseases. Nat Rev Genet 2005;6:287–98.

172. Wu Y, Zhu X, Chen J, Zhang X. EINVis: a visualization tool for analyzing and exploring genetic interactions in large-scale association studies. Genet Epidemiol 2013;37:675–85.

173. Wang X, O'Connell K, Jeon J, Song M, Hunter D, Hoffmeister M, et al. Combined effect of modifiable and non-modifiable risk factors for colorectal cancer risk in a pooled analysis of 11 population-based studies. BMJ Open Gastroenterol 2019;6: e000339.

174. Aschard H, Chen J, Cornelis MC, Chibnik LB, Karlson EW, Kraft P. Inclusion of gene-gene and gene-environment interactions unlikely to dramatically improve risk prediction for complex diseases. Am J Hum Genet 2012;90:962–72.

175. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet 2019;51:584–91.

176. Kim MS, Patel KP, Teng AK, Berens AJ, Lachance J. Genetic disease risks can be misestimated across global populations. Genome Biol 2018;19:179.

177. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, et al. Human demographic history impacts genetic risk prediction across diverse populations. Am J Hum Genet 2017;100:635–49.

178. Wang S, Qian F, Zheng Y, Ogundiran T, Ojengbede O, Zheng W, et al. Genetic variants demonstrating flip-flop phenomenon and breast cancer risk prediction among women of African ancestry. Breast Cancer Res Treat 2018;168:703–12.

179. Ho WK, Tan MM, Mavaddat N, Tai MC, Mariapun S, Li J, et al. European polygenic risk score for prediction of breast cancer shows similar performance in Asian women. Nat Commun 2020; 11:3833.

180. Aragam KG, Natarajan P. Polygenic scores to assess atherosclerotic cardiovascular disease risk: clinical perspectives and basic implications. Circ Res 2020;126:1159–77.

181. Peterson RE, Kuchenbaecker K, Walters RK, Chen CY, Popejoy AB, Periyasamy S, et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. Cell 2019;179:589–603.

182. Duncan L, Shen H, Gelaye B, Meijsen J, Ressler K, Feldman M, et al. Analysis of polygenic risk score usage and performance in diverse human populations. Nat Commun 2019;10:3328.

183. Grinde KE, Qi Q, Thornton TA, Liu S, Shadyab AH, Chan KHK, et al. Generalizing polygenic risk scores from Europeans to Hispanics/Latinos. Genet Epidemiol 2019;43:50–62.

184. Coram MA, Fang H, Candille SI, Assimes TL, Tang H. Leveraging multi-ethnic evidence for risk assessment of quantitative traits in minority populations. Am J Hum Genet 2017;101:218–26.

185. Marquez-Luna C, Loh PR, Price AL. Multiethnic polygenic risk scores improve risk prediction in diverse populations. Genet Epidemiol 2017;41:811–23.

186. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc 2020;15:2759–72.

187. Pashayan N, Morris S, Gilbert FJ, Pharoah PDP. Cost-effectiveness and benefit-to-harm ratio of risk-stratified screening for breast cancer: a life-table model. JAMA Oncol 2018;4:1504–10.

188. Burnett-Hartman AN, Newcomb PA, Peters U. Challenges with colorectal cancer family history assessment-motivation to translate polygenic risk scores into practice. Gastroenterology 2020;158:433–5.

189. Henneman L, Timmermans DR, Bouwman CM, Cornel MC, Meijers-Heijboer H. 'A low risk is still a risk': exploring women's attitudes towards genetic testing for breast cancer susceptibility in order to target disease prevention. Public Health Genoms 2011;14:238–47.